IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 13, 2019, accepted January 26, 2019, date of publication March 5, 2019, date of current version March 20, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2897737*

# Multispectral Heterogeneity Detection Based on Frame Accumulation and Deep Learning

**BAOJU ZHANG[1,2], CHENGCHENG ZHANG[1,2], GANG LI[3,4], LING LIN[3,4], CUIPING ZHANG[1,2], FENGJUAN WANG[1,2], AND WENRUI YAN[1,2]**

[1]Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
[2]School of Electronics and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
[3]State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin 300072, China
[4]Tianjin Key Laboratory of Biomedical Detecting Techniques and Instruments, Tianjin University, Tianjin 300072, China

Corresponding author: Baoju Zhang (wdxyzbj@163.com)

This work was supported in part by the Natural Youth Science Foundation of China under Grant 61401310, and in part by the Tianjin Science Foundation under Grant 18JCYBJC86400.

**ABSTRACT** Hyperspectral transmission imaging may provide a means for rapid screening of breast tumors, but tissue has a strong nature of scattering, thus causing a great difficulty in identifying heterogeneity. In this paper, a combination of frame accumulation and deep learning was proposed to detect heterogeneity, and we designed the simulation experiment of collecting phantom images. On the basis of frame accumulation preprocessing, the heterogeneous detection is performed on multispectral images by using faster regions with convolutional neural networks (R-CNN) features and a single shot multibox detector (SSD), two typical detection frameworks of deep learning. The results show that the mean average precision (mAP) of faster R-CNN and SSD reach 90.8% and 95.1%, respectively, when three classes (including background) are detected with the help of the dataset provided in this paper, and the mAP of two frameworks both reach 99.9% when two classes (including background) are detected. The detection efficiency of the SSD is higher than faster, SSD's detection speed can reach 50 fps, and the detection accuracy of the images after frame accumulation preprocessing is higher than that without frame accumulation processing. In summary, we validate the possibility of employing faster R-CNN and SSD to detect heterogeneity in multispectral images based on frame accumulation that improves image grayscale resolution, and it has a certain degree of reference significance for the application of deep learning in multispectral image detection.

**INDEX TERMS** Deep learning, frame accumulation, heterogeneity detection, multispectral image.

## I. INTRODUCTION

At present, breast cancer has become a major killer of women [1], and its early detection is beneficial to the improvement of the cure rate. However, there are many shortcomings in existing clinical examination, such as high price and long detection time Therefore, it is necessary to conduct research on the self-test methods of early-stage breast tumors. In recent years, scholars have explored hyperspectral transmission tissue imaging and its treatment and screening for breast tumors [2]–[6]. Walter *et al.* [7] used breast spectroscopy to predict the density of mammograms by placing the light source and detector in a fixed position in four-sized breast-adapting cups to predict breast lesions. Yang *et al.* [8] used solid and liquid as the phantom of biological tissues and heterogeneous tissues respectively to obtain the hyperspectral image by transillumination method and assess the spatial information of the lesion. However, the strong scattering of the tissue leads to weak signal and low signal-to-noise ratio (SNR) of the transmitted image, which is not conducive to edge detection and tissue classification, bringing great difficulties in identifying heterogeneity In terms of improving the quality of transmission images, wavelet transform filtering method, space-time domain joint filtering method and some classical image denoising methods (assuming noise obeys Gaussian distribution, Poisson distribution, etc.) [9], [10] have been proposed in recent years but the filtering method often causes the loss of the image's edge details while denoising. The frame accumulation technique that has been successfully applied to various low-light-level

The associate editor coordinating the review of this manuscript and approving it for publication was Qilian Liang.

image detection devices is one of the most effective methods for enhancing weak transmission image signals. Li *et al.* [11] and Yu *et al.* [12] greatly enhanced the signal-to-noise ratio (SNR) of the low-light transmission image by combining the frame accumulation and the shaping signal technology, and improve the detection sensitivity of the transmission image. In heterogeneous detection, many excellent algorithms have emerged in the field of object detection based on deep learning [13], [14], such as R-CNN [15], Fast R-CNN [16], Faster R-CNN [17], SSD [18], YOLO [19] and so on, among which Faster RCNN and SSD are typical representatives and have achieved excellent detection results under the VOC2007 data set. However, no predecessors have combined the two algorithms and applied in the detection of multispectral heterogeneity.

Therefore, we, for the first time, propose a method combining the frame accumulation technique and deep learning to detect heterogeneous tissues The experiment results show that, under the data set of this experiment, Faster R-CNN and SSD achieve 90.8% and 95.1% mAP [20] respectively while detecting 3 classes (including background), and both achieve 99.9% mAP while detecting 2 classes (including background) based on the frame accumulation improving the multi-spectral gray-scale resolution. The image detection accuracy is higher after preprocessing by frame accumulation than that of the image without frame accumulation, and the detection efficiency of SSD is greater than Faster, reaching a testing speed of 50fps which means that it takes only 0.02s to detect a 502 $\times$ 362 image on a NVIDIA Tesla P100 GPU, and almost a real-time detection is achieved. In a word, we verify the feasibility of applying Faster RCNN and SSD to detect two-dimensional multispectral images based on the frame accumulation preprocessing, which possibly alternates a testing method of heterogeneity detection with breast hyperspectral or multispectral transmission images.

## II. RELATED TECHNOLOGY

### A. FRAME ACCUMULATION TECHNIQUE

The facts that the light intensity may be affected by the environment and individuals and the biological tissue has comparatively strong scattering and absorption characteristics in the process of self-test of the human body are taken into consideration, which thus inevitably leads to weak signal and insufficient image clarity. The weak signal has two implications: one is that the absolute value of the signal intensity is low, and the other is that the SNR and resolution of the signal are low. The frame accumulation technique that has been successfully applied to various low-light-level image detection devices is one of the most effective methods for detecting weak image signals. Therefore, we use frame accumulation technique to improve the SNR of images and gray-scale resolution to obtain high-precision multispectral images.

Gray level is the core of image accuracy and heterogeneity detection sensitivity. The higher the gray level resolution of the image is, the richer the image information is, the better the classification and analysis of the tissue will be. Low gray resolution becomes an obstacle to tissue classification and spatial information extraction, but the grayscale resolution can be improved to some extent by frame accumulation [11]. In the image processing algorithm, multi-frame accumulation is to add the gray values of the corresponding pixels of two images or multiple frames of images at different times to obtain their time-averaged images, which can multiply the SNR of the image.

### B. OBJECT DETECTIONS

Currently, the object detection algorithm based on deep learning is generally divided into two categories: the R-CNN series based on regional proposals and the YOLO and SSD series without regional proposals, and Faster RCNN and SSD are typical representatives of these algorithms separately.

Faster R-CNN abandons the Selective Search [20] algorithm in this type of algorithm but introduces the RPN network, and adopts regional proposals to make the classification and regression share the convolution feature, thus further accelerating the detection process

SSD was an object detection framework proposed in 2016. It has been one of the main detection frameworks so far. It has obvious speed advantage over YOLO, and remarkable mAP advantage over Faster R-CNN. SSD is a method of detecting objects in an image which employs a single deep neural network [18] and totally eliminates the resampling phase of candidate frames (A region candidate box indicating that n candidate frames are extracted from the possible regions in the image and then classified training) and their subsequent pixels or features in this type of region candidate and packages all computations in a single network, which makes SSD easy to train and integrate directly into systems that are needed to detect components. Experiment results from the PASCALVOC, MS COCO and ILSVRC data sets confirm that as for the accuracy, SSD is almost the same as that of regional proposal method, but much faster. SSD, even with a smaller input image size, still enjoys a great accuracy compared with other single-stage methods. For 300 $\times$ 300 input, SSD achieves 74.3% mAP on VOC2007 test at 59 FPS on a Nvidia Titan X outperforming the Faster R-CNN model [17]. Moreover, SSD can also directly detect the video without decomposing it into frames, which greatly simplifies the steps of the pre-image preprocessing and improves the efficiency of the entire detection system.

## III. EXPERIMENT

According to the characteristics of breast tissue, we designed a simulation experiment of collecting the phantom image and found a better pretreatment method for heterogeneity detection based on the frame accumulation (Here, ''better'' means that the cooperation with the detection of the heterogeneous body is better) Then we detected heterogeneity of multispectral images using the two typical object detection frameworks of deep learning – Faster R-CNN and SSD. A series of procedures were taken: fine-tuning the two object

detection frameworks, employing a very deep VGG16 network [21], and adopting both enhanced and unenhanced edges of multi-spectral abstract images as training sample sets. Then each set was trained and tested in the Faster RCNN and SSD frameworks. Finally, the heterogeneity detection was performed, the heterogeneous tissues were marked, and then its spatial position information was obtained.

### A. EXPERIMENTAL DEVICE

Figure 1 depicts a schematic diagram of the experimental equipment The device consists of the following components: LED light source (0.5W), imitation [2], mobile phone (model: Huawei mate9, frame rate: 59 fps, image resolution: 1080 × 1920), a computer (connected to mobile phone) for images handling, shading cloth. Phantom is composed of a PMMA (polymethyl methacrylate) cuboid container with a transmittance of 96% (the transmittance of the breast is higher than that of other human tissues), which contains a mixed solution of milk and water. We suspend pork and carrot cubes in the solution representing heterogeneity tissues in terms of optical properties of different materials under different wavelengths

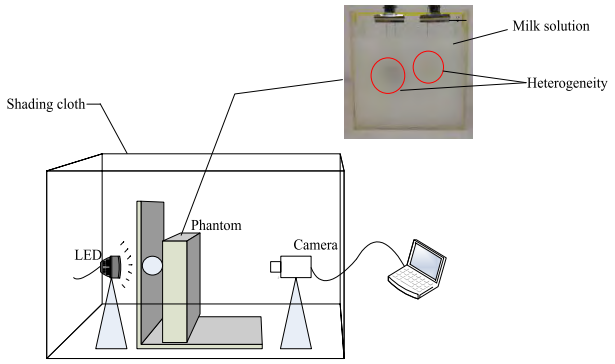


**FIGURE 1.** A schematic diagram of the experimental equipment, and the inset is a photo of an actual phantom example.

### B. IMAGE ACQUISITION

Based on the experiment setups, the original multispectral images are obtained: milk solution + heterogeneity image.

The specific experiment procedures are as follows:

(1) Adjust and fix the distance between the light source and the experimental phantom and the distance between the mobile phone and the experimental phantom, and set up the blackout cloth. Turn on the light source and the phone camera to record the video of this phantom.

(2) The experiment was carried out in groups: 12 minutes of video images for each group and a total of $k = 36$ groups. The size, thickness, shape and number of heterogeneous tissue in each group are different.

(3) Extract images from each video, and get rid of the images with obvious errors in each group, then get $k = 36$ kinds, a total of $n = 150,0000$ frames of original multispectral phantom images $x_i^q$ ($q = 1, 2, \ldots, k$; $i = 0, 1, 2, \ldots, n-1$) Figure 2 (a) is the original phantom image in a certain group

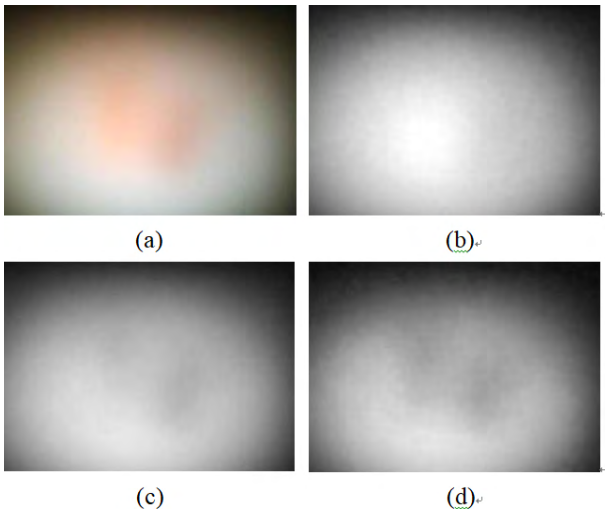


**FIGURE 2.** The original phantom image and its three single-channel images. (a) The original phantom image $x_i^q$. (b)–(d) Show the R, G, and B single-channel grayscale image $x_{i,1}^q, x_{i,2}^q, x_{i,3}^q$, respectively.

(4) Each group of images in (3) is separated from RGB, and to obtain three single-channel images, and finally $k = 36$ kinds are achieved, and a total of $m = 3n$ single-channel grayscale images is gained

### C. IMAGE PREPROCESSING

We found a better pretreatment method with the heterogeneity detection on account of the frame accumulation mentioned in this paper.

(1) The original image of size 1080 × 1920 is reasonably cropped to 500 × 360 without affecting the accuracy and quality of the algorithm. For each group of single-channel images $x_{i,j}^q$, each N frames is averagely frame-accumulated. That is, for each group of images $q = 1, 2, \ldots, k$ we have:

$$x_t^q = \begin{cases} \sum_{i}^{i+N-1} x_{i,1}^q = x_{i,1}^q + x_{i+1,1}^q \\ \quad + x_{i+2,1}^q + \cdots x_{i+N-1,1}^q \\ \sum_{i}^{i+N-1} x_{i,2}^q = x_{i,2}^q + x_{i+1,2}^q \\ \quad + x_{i+2,2}^q + \cdots x_{i+N-1,2}^q \\ \sum_{i}^{i+N-1} x_{i,3}^q = x_{i,3}^q + x_{i+1,3}^q \\ \quad + x_{i+2,3}^q + \cdots x_{i+N-1,3}^q \end{cases}$$
$$0 \le i < n-1, \quad N > 1, \; t = 1, 2, \ldots, \frac{n}{N}. \quad (1)$$

average results after accumulation:

$$\bar{x}_t^q = \frac{1}{N} \sum_{i}^{i+N-1} x_{i,j}^q, \quad t = 1, 2, \ldots, 15000; \quad q = 1, 2, \ldots, k \quad (2)$$

In the experiment, N is taken the value of 100, then get the 15,000 mean frame-accumulated images $\bar{x}_{t,j}^{q} = x_{t,1}, x_{t,2}, x_{t,3} (q = 1, 2, \ldots, k; t = 1, 2, \ldots, 15000, j = 1, 2, 3)$ of the original multispectral single-channel images according to formula (1) (2). Then it is subjected to a series of filtering and denoising operations.

(2) The single-channel grayscale images $\bar{x}_{t,j}^{q} (q = 1, 2, \ldots, k; t = 1, 2, \ldots, 15000, j = 1, 2, 3)$ obtained through averaging the frame-accumulated images are reconstructed to 15,000 color images $x_r$ $(r = 1, 2, \ldots, 15000)$ totally, and then edge-enhanced processing is performed to acquire edge-enhanced images $x_h$ $(h = 1, 2, \ldots, 15000)$.

### D. HETEROGENEITY DETECTION

#### 1) DATASET PRODUCTION

The images $x_r$ $(r = 1, 2, \ldots, 15000)$ is made into dataset-a, and their contents are divided into three classes: background, raw pork and carrot. The images $x_h$ $(h = 1, 2, \ldots, 15000)$ are processed into dataset-b, and their contents are classified into two classes: background and heterogeneity (raw pork and carrot are considered heterogeneity). Both a and b datasets are randomly divided into training set, validating set and testing set, and due to the small data class in our paper, the ratio is set to 6:2:2 according to the traditional division in the machine learning field.

#### 2) TRAINING FASTER R-CNN AND SSD

Fine-tuning [22]: we modified the training parameters of the two frameworks, such as learning rate, step size, and the number of iterations, and debug for the dataset of the experiment And we used ImageNet to pre-train the network model. Samples of the heterogeneous tissues + milk solution images of dataset a and b were imputed into the network for training. In order to find the optimal model, a total of two data sets a and b were tested and trained four times. The learning rate and the number of iterations were 0.001 and 80k, 0.0001 and 80k, 0.0001 and 40k, 0.0001 and 20k. Finally, a special detection model for the phantom image in the paper was obtained.

#### 3) TESTING

We prepared 3 kinds of test data for comparing: ①: testing set image data within the data set of the experiment. ②: a small amount of preprocessed image data excluded from the data set of the experiment. ③: a small amount of unprocessed image data beyond the data set of the experiment. Three kinds of data were tested sequentially to obtain the position, score, category of the heterogeneity.

## IV. RESULTS AND ANALYSIS

### A. ANALYSIS OF FRAME ACCUMULATION PREPROCESSING RESULTS

After the multispectral images are processed by frame accumulation, the SNR is significantly enhanced, the grayscale level is stretched, and the grayscale resolution is improved, on the basis of which heterogeneous body detection experiment based on deep learning has achieved excellent results.

Frame accumulation significantly improves the quality of the phantom image. It is discovered through calculation that the peak signal-to-noise ratio (PSNR) of the image after frame accumulation is 57.32dB in the experiment, which is in accordance with the theoretical derivation in section II. A of the paper. Since the existing display device has only 256 gray levels, it is impossible to display images of higher gray levels. To show the effect of frame accumulation, we normalized the image and then stretched it to 256 grayscale The original single-channel images their N frame accumulated ones and the reconstructed color images after averaging frame-accumulated images are shown in Figure 3.

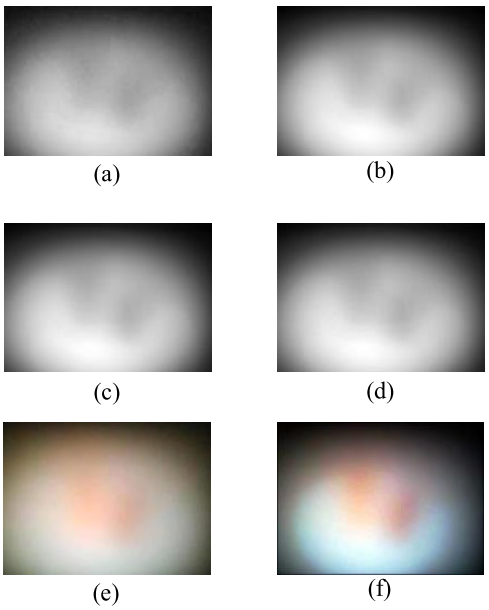


**FIGURE 3.** Comparison of the original images and their N frame accumulated ones. (a) The B channel image of the original image $x_i^q$. (b) The B channel image $x_t^q$ of $N$=100 frames of accumulation. (c) The B channel image $x_t^q$ of $N$=1000 frames of accumulation. (d) The mean image $\bar{x}_t^q$ of $N$=100 frames accumulated image. (e) The original color image $x_i^q$. (f) The reconstructed color images after averaging frame-accumulated images.

After filtering and frame accumulation averaging, the image quality of this experiment is evaluated by using the quality evaluation method without reference image [23]. Approximating the 3000-frame accumulated images as a true-value image, it is found that the mean signal-to-noise ratio (MSNR) of the single-channel grayscale image after frame accumulation is significantly improved.

### B. HETEROGENEITY TESTING RESULTS

Fine-tuning the Fatser and SSD frameworks based on frame accumulating improved multispectral quality.

The two datasets of a and b were trained in multiple frames under two different frameworks with multiple learning rates and iterations to obtain the best detection model. In the object

**TABLE 1.** The MSNRs of the original single-channel image and the N-frames accumulated image.

| Image | MSNR/dB |
|---|---|
| Original image | 38.5706 |
| 100 frames of accumulated images | 47.6169 |
| 1000 frames of accumulated images | **59.0384** |

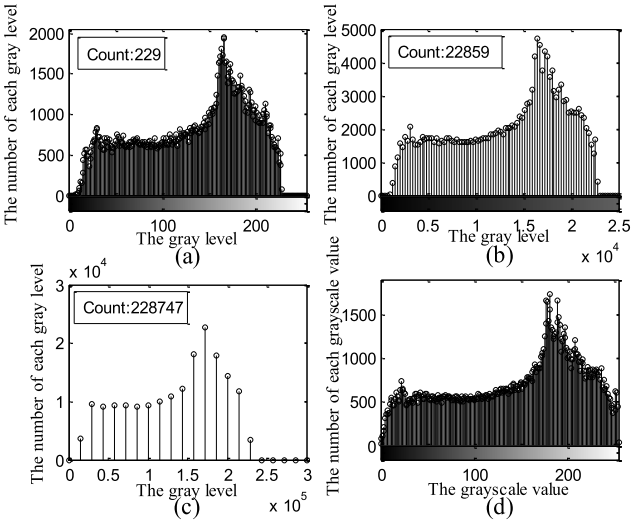


**FIGURE 4.** Histogram of the cumulative gray of the original image and the multi-frame accumulated image, the histogram of the mean image of the 100-frames accumulation and (a)–(d) are the gray histograms of (a)–(d) of Figure 3, respectively. (a) The histogram of the original image. (b) The cumulative histogram of the accumulated 100 frames. (c) The cumulative histogram of the accumulated 100 frames. (d) The histogram of $\bar{x}_t^q$ that 100-frames accumulated mean image, and the x-axis of (d) is the gray level of $\bar{x}_t^q$, the y-axis is the number of each gray level. And the x-axis of (a), (b), and (c) is the cumulative gray value, the y-axis of (a)–(c) is the number of each cumulative gray level. Count represents the number of gray levels of each accumulated image.

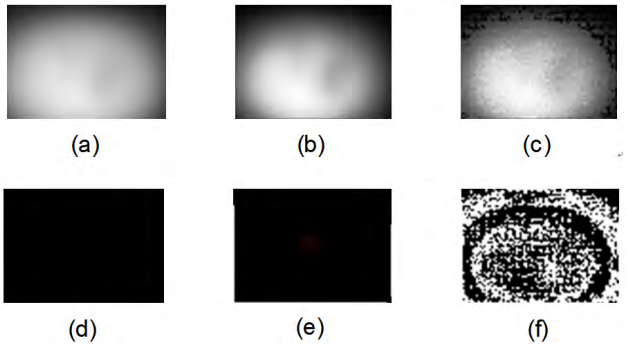


**FIGURE 5.** The comparison of the grayscale image of original image $x_i^q$, the reconstructed image $x_r$ and the edge-enhanced image $x_h$ and its edge detection image. (a)– (c) sequentially show the grayscale image of original image, the reconstructed image after preprocessing by frame accumulation, and the edge-enhanced image. (d)–(f) are edge detection images of (a)–(c), respectively.

**TABLE 2.** The mAPs of models that Faster R-CNN and SSD frameworks trained in different training parameters for two datasets of a and b. Base_lr: the base learning rate.

| Method | Base_lr | Iteration | mAP/% |
|---|---|---|---|
| Faster (Dataset a) | $10^{-3}$ | 80k | 90.8 |
| Faster (Dataset a) | $10^{-4}$ | 20k | 90.8 |
| | | 40k | 90.7 |
| | | 80k | 90.8 |
| SSD (Dataset a) | $10^{-4}$ | 20k | 90.8 |
| | | 40k | 90.8 |
| | | 80k | 95.1 |
| Faster (Dataset b) | $10^{-3}$ | 80k | 99.9 |
| Faster (Dataset b) | $10^{-4}$ | 20k | 99.9 |
| | | 40k | 99.9 |
| | | 80k | 99.9 |
| SSD (Dataset b) | $10^{-4}$ | 80k | 99.9 |

detection, Average Precision (AP) is the average accuracy of each class, and mAP [20] is the average value of all APs. For all APs of each class tested, the average value is achieved, thus mAP is obtained. The mAP is used as an indicator for measuring the accuracy of detection in the object detection.

The total number of detection classes is represented by $Q_R$, then:

$$mAP = \frac{1}{|Q_R|} \sum_{q \in Q_R} AP(q) \tag{3}$$

TABLE 2 shows the mAPs of the two frameworks in different training parameters.

The final 4 models selected from TABLE 2 are trained in $10^{-4}$ learning rate and iterating 80k times: Model1: F + a, Model2: S + a, Model3: F + b, Model4: S + b. Explanation for the above models: F + a: data set a is trained on Faster R-CNN; S + a: data set a is trained on SSD; F + b: data set b is trained on Faster R-CNN; S+b: data set b is trained on SSD.

We use these models as references for the final detection of heterogeneity.

It is obtained from TABLE 3: in the experiment, the mAP of Faster R-CNN trained by the phantom image in the paper is slightly lower than SSD, and the detection speed

**TABLE 3.** The mAPs and detection rates of the Faster R-CNN and SSD on the test set of datasets-a and datasets-b respective.

| Method | Base_lr | Iteration | mAP/% | rate |
|---|---|---|---|---|
| Faster (Dataset a) | $10^{-3}$ | 80k | 90.8 | 11fps |
| Faster (Dataset b) | $10^{-3}$ | 80k | 99.9 | 14fps |
| SSD (Dataset a) | $10^{-4}$ | 80k | 95.1 | 25fps |
| SSD (Dataset b) | $10^{-4}$ | 80k | 99.9 | **50fps** |

**TABLE 4.** Model1's detection accuracy for ①, ② and ③ of test data.

| Test data | Accuracy/% |
|---|---|
| model1+① | **85.7** |
| model1+② | 75.0 |
| model1+③ | 66.7 |

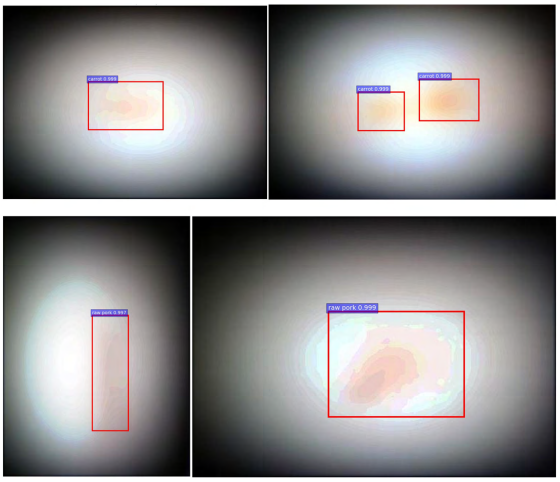

**FIGURE 6.** The detection examples on Model1 that the image contents are divided into three classes (including background).

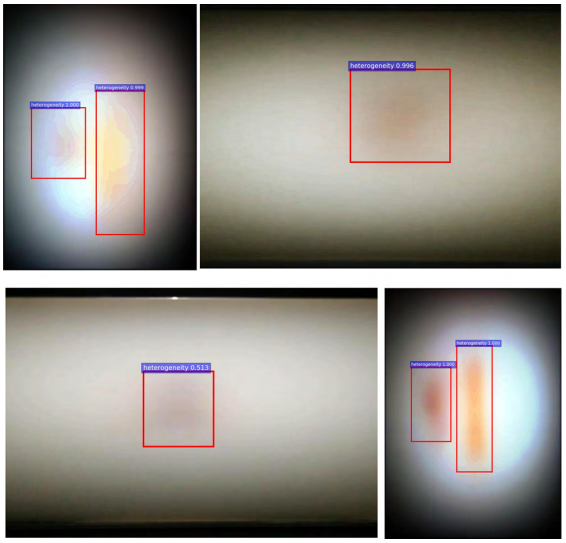

**FIGURE 7.** The detection examples on Model3 that the image contents are divided into two classes (including background).

is also slower than SSD. Table 4 shows the detection accuracy of Model1 for each of the three kinds of test data. ①: test set image data within the data set of experiment, total of 3000 images. ②: a small amount of preprocessed image data excluded from the data set of experiment, total of 96 image. ③: a small amount of unprocessed image data not within the data set of experiment, total of 96 images.

It can be seen from Table 4 that the detection accuracy of the three kinds of test data is ① > ② > ③, which indicates that the SNR and resolution of the image have an influence on the detection accuracy. It also indicates that the frame accumulation technique improved the accuracy of detection to a certain extent, which is consistent with our expectation in the previous part of the paper.

Figure 6 and Figure 7 shows some examples of 3 kinds of test data for the model trained by datasets of a and b on Faster R-CNN.

Figure 8 and Figure 9 shows some examples of 3 kinds of test data for the model trained on the SSD.

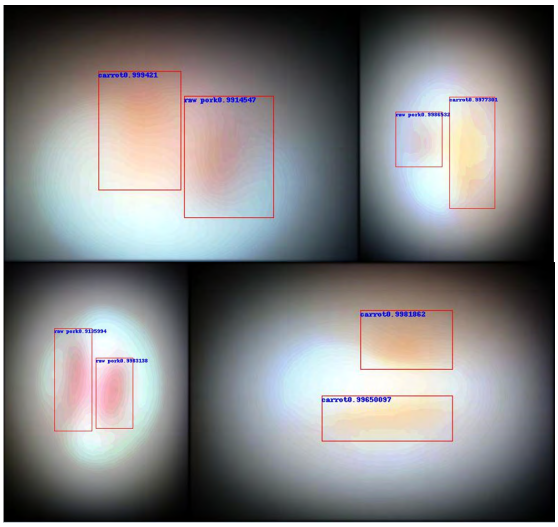

**FIGURE 8.** The detection examples on Model2 that the image contents are divided into three classes (including background).

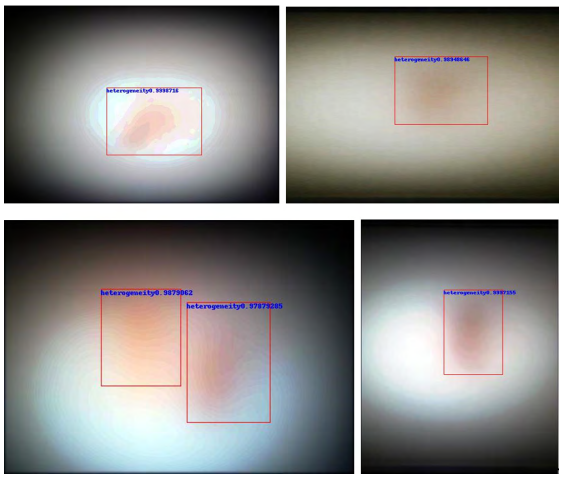

**FIGURE 9.** The detection examples on Model4 that the image contents are divided into two classes (including background).

Analysis of testing results: regarding mAP as an indicator, and comparing the experiment data of Faster and SSD, it can be found that the mAP of the SSD is slightly higher than the Faster R-CNN in the same learning rate and number of iterations, and in terms of detection speed, the SSD is

also slightly faster. It is found that the detection accuracy of the data subjected to the frame accumulation processing is higher than that of the data without such processing by comparing the accuracy of the three kinds of test data, which is in accordance with the theoretical analysis: the image processed by the frame accumulation method used in the paper enjoys higher SNR, sharper resolution, and more prominent heterogeneity.

## V. CONCLUSION

In the paper, according to the characteristics of breast tissue, we designed the multispectral phantom image acquisition experiment, and then combined frame accumulation and deep learning to study the heterogeneity detection of multispectral images. In terms of image preprocessing, a better pretreatment method with the heterogeneity detection based on the frame accumulation is found. Then we use Faster R-CNN and SSD object detection framework concerning deep learning, the high-quality dataset is trained as a sample. Finally, an excellent testing result is obtained: the values of mAP of Faster R-CNN and SSD reach 90.8% and 95.1% respectively when 3 classes (including background) are detected under the dataset in the paper, and the mAP of two frameworks both reach 99.9% when 2 classes are detected (including background). The detection efficiency of SSD is higher than Faster and detection speed of SSD can reach 50fps, and more importantly, the accuracy of the images after frame accumulation preprocessing is higher than that without frame accumulation processing. In summary, we validate the possibility of using Faster RCNN and SSD to detect heterogeneity in multispectral images based on frame accumulation that improves image grayscale resolution, and it has a certain reference significance for the application of deep learning in multispectral image detection.

## REFERENCES


[1] L. Fan *et al.*, "Breast cancer in China," *Lancet Oncol.*, vol. 15, no. 7, pp. e279–e289, Jun. 2014.

[2] G. Lu and B. Fei, "Medical hyperspectral imaging: A review," *J. Biomed. Opt.*, vol. 19, no. 1, Jan. 2014, Art. no. 10901. doi: 10.1117/1.JBO.19.1.010901.

[3] K. Van de Vijver, E. Kho, L. de Boer, H. Sterenborg, and T. Ruers "Hyperspectral optical imaging for intraoperative margin assessment during breast cancer surgery," *Virchows Archiv*, vol. 473, no. 1, p. S200, Sep. 2018.

[4] H. Jiang, S.-C. Yoon, H. Zhuang, W. Wang, K. C. Lawrence, and Y. Yang, "Tenderness classification of fresh broiler breast fillets using visible and near-infrared hyperspectral imaging," *Meat Sci.*, vol. 139, pp. 82–90, May 2018. doi: 10.1016/j.meatsci.2018.01.013.

[5] B. J. Tromberg *et al.*, "Non-invasive *in vivo* characterization of breast tumors using photon migration spectroscopy," *Neoplasia*, vol. 2, nos. 1–2, pp. 26–40, Jan. 2000.

[6] P. Taroni, A. Pifferi, E. Salvagnini, L. Spinelli, A. Torricelli, and R. Cubeddu "Seven-wavelength time-resolved optical mammography extending beyond 1000 nm for breast collagen quantification," *Opt. Express*, vol. 17, no. 18, pp. 15932–15946, Aug. 2009.

[7] E. J. Walter, J. A. Knight, and L. Lilge, "A multi-wavelength, laser-based optical spectroscopy device for breast density and breast cancer risk prescreening," *J. Biophoton.*, vol. 10, no. 4, pp. 565–576, Apr. 2017.

[8] X. Yang, G. Li, and L. Lin, "Assessment of spatial information for hyperspectral imaging of lesion," *Proc. SPIE*, vol. 10024, Oct. 2016, Art. no. 100242O. doi: 10.1117/12.2245874.

[9] P. Roudot, C. Kervrann, J. Boulanger, and F. Waharte, "Noise modeling for intensified camera in fluorescence imaging: application to image denoising," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, Apr. 2013, pp. 600–603. doi: 10.1109/ISBI.2013.6556546.

[10] J. Jin, B. Yang, K. Liang, and X. Wang, "General image denoising framework based on compressive sensing theory," *Comput. Graph.*, vol. 38, pp. 382–391, Feb. 2014.

[11] G. Li, H. Tang, D. Kim, J. Gao, and L. Lin, "Employment of frame accumulation and shaped function for upgrading low-light-level image detection sensitivity," *Opt. Lett.*, vol. 37, no. 8, pp. 1361–1363, Apr. 2012.

[12] J. Yu, G. Li, S. Wang, and L. Lin "Employment of the appropriate range of sawtooth-shaped-function illumination intensity to improve the image quality," *OPTIK*, vol. 175, pp. 189–196, Dec. 2018.

[13] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *MIT Press J.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[14] L. Yann, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[15] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. CVPR*, Jun. 2014, pp. 580–587.

[16] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.

[17] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[18] W. Liu *et al.*, *SSD-Single Shot MultiBox Detector*. Cham, Switzerland: Springer, 2016, pp. 21–37.

[19] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788.

[20] J. R. R. Uijlings, E. A. Van De Sande Koen, G. Theo, and W. M. S. Arnold, "Selective search for object recognition," *Int. J. Comput. Vis.*, vol. 104, no. 2, pp. 154–171, Sep. 2013.

[21] K. Simonyan and A. Zisserman. (2014). "Very deep convolutional networks for large-scale image recognition." [Online]. Available: https://arxiv.org/abs/1409.1556

[22] Z. Zhou, J. Shin, L. Zhang, S. Gurudu, M. Gotway, and J. Liang, "Fine-tuning convolutional neural networks for biomedical image analysis: Actively and incrementally," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4761–4772. doi: 10.1109/CVPR.2017.506.

[23] J. Yu, G. Li, S. Wang, and L. Lin, "Image quality assessment metric for frame accumulated image," *Rev. Sci. Instrum.*, vol. 89, no. 1, Jan. 2018, Art. no. 013703. doi: 10.1063/1.5020715.


**BAOJU ZHANG** received the B.S. and M.S. degrees from Tianjin Normal University, in 1990 and 1993, respectively, and the Ph.D. degree from Tianjin University, in 2002. She is currently a Professor with the College of Electronic and Communication Engineering, Tianjin Normal University. Her main research interests include compressive sensing, signal processing, audio and video processing, data stream clustering, image processing, image recognition, and object detection.

**CHENGCHENG ZHANG** was born in Heze, China, in 1994. She received the B.S. degree from Ludong University, in 2016. She is currently pursuing the M.S. degree in communication engineering with Tianjin Normal University. Her research interests include image processing, deep learning and object detection, multispectral image detection, and hyperspectral transmission imaging.

**GANG LI** received the B.Sc. degree from the Hefei University of Technology, in 1982, and the M.Sc. and Ph.D. degrees from Tianjin University, in 1987 and 1998, respectively, where he is currently a Professor and the Supervisor for Ph.D. candidates with the School of Precision Instrument and Opto-Electronics Engineering. His main research fields and specialties are signal detection and analysis, precision medical instrument, biological image processing, and tissue image detection.

**LING LIN** received the B.S. degree in precision measurement instruments from Hefei Polytechnic University, in 1982, the M.S. degree in measurement technology and instruments from Tianjin University, in 1987, and the Ph.D. degree in biomedical engineering from Hokkaido University, in 1999. She is currently a Professor with the School of Precision Instrument and Opto-Electronics Engineering, Tianjin University. Her research interests include signal detecting and processing, biological image processing, tissue image detection, sensor engineering, and microcomputer application.

**CUIPING ZHANG** received the M.S. degrees from Tianjin Normal University, Tianjin, China, where she is currently an Assistant Experimenter with the College of Electronic and Communication Engineering. Her main research interests include wireless sensor networks, image processing, and object detection.

**FENGJUAN WANG** is currently pursuing the master's degree with Tianjin Normal University, Tianjin, China. Her research interests include image processing, object detection, deep learning, and machine vision.

**WENRUI YAN** is currently pursuing the master's degree with Tianjin Normal University, Tianjin, China. Her research interests include image processing, signal processing, and deep learning.